

SEGMENTATION METADATA FOR AUDIO-VISUAL CONTENT

BACKGROUND OF THE INVENTION

The present invention relates to descriptions of audio-visual material.

5 Digital audiovisual material is becoming increasingly available to users through digital TV broadcast, digital video cameras, digital video discs, and personal computer based access to multimedia on the Internet or other network. In addition, persistent large-volume storage and non-linear access to audiovisual content is becoming available in consumer devices. Consequently, there is a need for rapid navigation and searching capabilities to enable users to efficiently discover and consume the contents of
10 audiovisual material.

The extensive proliferation of audio-visual material available to users has the potential to overwhelm the consumer and lead to frustration at the inability to search and view content in an efficient manner. Viewing summaries of the content allows the consumer to skip irrelevant content and view the desired content quickly and easily.
15 Further, multiple different summaries, if available, may provide the user with alternative views of a particular program that the user could choose from depending on personal preferences or usage conditions.

Limited summary selection capabilities are appearing more frequently in
20 current technologies, such as the digital video disk (DVD). DVD movies normally provide "scene selections" or "chapter selections" that have a visual array of thumbnails and textual titles associated with each scene. This permits the user to click on the thumbnail of the desired scene, jump to that scene, and start playback. Playback typically

5 continues until the end of the movie, unless the user makes another selection. While somewhat limited, these features provide the capability to index for the purpose of jumping to an arbitrary position and continue playback from that position.

Referring to FIG. 1, a system 10 may be used for the presentation of video, audio, or audiovisual information to a user. The information may be presented to the user using any suitable medium, such as for example, the Internet, a computer network, the radio, a personal computer, and a television. The user or the user's agent interacts with the system 10 to receive the information in a desirable manner and to define preferences as to what type of information is obtained. The term user is intended to refer to the recipient of the information, which may be for example, a person, a machine, or a software program operating on a machine.

To define these interactions, a set of description schemes containing data describing the content of the material may be used. User preferences 12 may be used in several different areas to maximize both the user's enjoyment and the system functionality. The preferences describing the topics and subject matter of interest to the user is used in both searching for and navigating the audiovisual programs 14. These two sets of data, the user preferences 12 and program descriptions 14, are correlated in the filtering and search engine 16 to identify the preferred programs.

The programs identified by the filtering and search engine 16 are then forwarded to a browsing module 18 along with the user's browsing preferences. Another output of the filtering and search engine 16 are preferred programs that the user has designated for storage. These are stored in the storage module 20. The programs selected by the user with the browsing module 18 are then sent to a display 22. The user may

5 utilize multimedia title descriptions of preferred programs to navigate among the
programs that the user wants to consume. Once a program is selected, a summary
description of that particular program is correlated with user's browsing preferences to
offer the user a preferred summary.

10 The display 22 receives the programs and displays them in accordance with
the user's device preferences as to the operation of the display. User's device preferences
may include, for example, device settings such as volume setting that may vary with the
genre of the program that is being consumed. The display and user's interaction with the
display, such as stopping a program before its end and consuming certain types of programs
with certain device settings, also provides information in a manner analogous to a feedback
15 loop to update and log the usage history 24. The usage history 24 may be mapped against
the preferences by a mapping module 26. This information is then used in conjunction with
user inputs by the user preference module 12.

20 These user preferences may be useful in many contexts, not just an
audiovisual presentation system. User preferences and usage history may be transmitted
to the provider of audiovisual programs 14 to receive selected programming or directly
receive program segments that are preferred by the user. In the latter case, user
preferences may be correlated with summary descriptions at the provider side to select
and directly deliver summarized audiovisual programs to the user. The preferences may
also be transferred to a "smart card" 28 or similar, portable storage and ultimately
25 transferred to another system by the user.

However, it is noted that a framework for the description of the individual
description schemes at the user, program, or device level are needed. As illustrated in

4005369.01300

5 FIG. 5 illustrates permitted grouping.

FIG. 6 illustrates alternative groups.

FIG. 7 illustrates a group and related segments.

FIG. 8 illustrates a table of contents presentation.

FIG. 9 illustrates a virtual program.

10 FIG. 10 illustrates a temporal time base for video.

FIG. 11 illustrates entity relationships.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

The present inventors came to the realization that the previously existing searching description scheme (segment description scheme in MPEG-7) and the navigation description scheme (summary description scheme in MPEG-7), as described in MPEG-7, are inconsistent with one another if it is desirable to navigate portions of a video and simultaneously obtain information regarding the content of those portions. In particular, MPEG-7 does not provide sufficient syntax structures to physically or logically link the segments identified by the navigation description scheme with the video identified by the searching description scheme. Referring to FIG. 3, a modified description scheme syntax incorporates both information related to the searching description scheme and the navigation description scheme (e.g., segmentation description scheme), both within a single description scheme. Moreover, the incorporation of both the searching description scheme and the navigation description scheme into a single description scheme eliminates some redundant information. Within an XML syntax

15

20

25

5 structure both the searching and the navigational descriptions are associated with the same segment definition (e.g., specifies physical location of segment within the video).

The description scheme structure within MPEG-7 permits a hierarchical nesting structure for segment descriptions of the video and descriptions for groups of segments of the video. The permitted hierarchical nesting structure is very flexible and permits nearly any desirable interrelationship to be defined. Referring to FIG. 4, a sample hierarchical nesting structure is illustrated. While such a hierarchical nesting structure is flexible, it turns out to be extremely difficult determining the order to present the segments of video within such a structure. Within a particular level it is difficult to determine which segments to present first and is further complicated by determining which segment to present within the different hierarchical levels. Moreover, the parsing of such a structure is complex which reduces the computational efficiency of the system.

To overcome the non-deterministic nature of the hierarchical structure the preferred system imposes at least one or more of the following restrictions. A segment group may reference either other segments or other segment groups, but not both, as illustrated in FIG. 5. A segment group references other segments or other segment groups, but does not contain segments itself. In some cases a segment group may reference a single segment or otherwise contain a single segment, in which case it is considered a segment. In this manner, any particular segment group will either reference a set of segments which may be readily ordered, or one or more segment groups.

Segments may not reference other segments or segment groups.

With the set of permissible interconnections limited a set of rules is useful in order to permit the user to view the available content in an organized manner or

otherwise select a set of segments for presentation in a particular order. In this manner the playback, navigation, and presentation order may be unambiguously defined.

Therefore different systems will interpret the segmentation data the same. One type of organizational technique is to define segments type “alternativeGroups”, which may not contain segments and shall only contain subgroups. The user may select one of the

groups from the set of groups at the same hierarchical level originating from the same parent group. Referring to FIG. 6, for example, the user may select Group 1 or Group 2.

If Group 1 is selected, then the user may select Group 3 or Group 4. Group 3 consists of one segment. If Group 4 is selected, then the set of segments referenced by Group 4 are made available to the user. If the user selects Group 4 the segments defined within Group

4 may be presented, as illustrated in FIG. 7. The order of presentation of the segments defined by a Group may be presented to the user in an order defined by the order listed within the Group, such as the first in the list is first played or the last in the list is first played.

Another type of organizational technique is to define a set of segments of type “tableOfContents”, which presents the groups and segments defined therein in an ordered manner such that the hierarchical order may be observed. For example, a portion of the ordered groups shown in FIG. 6 may be organized as a tableOfContents, as shown in FIG. 8. The tableOfContents representation facilitates the user being able to select portions of video content that may be of particular interest to the user in a convenient manner.

Portions of the hierarchical structure of segment groups may be designated as “alternativeGroups” and other portions of the hierarchical structure may be designated

5 as "tableOfContents". Preferably, the two different designations of the hierarchical structure are non-overlapping, but may be overlapping, if desired. These designations are preferably not directly associated with segments.

Existing video summarization systems provide segmentation data for each video and permit the selective viewing of each video according to the segmentation data.

10 While beneficial, the present inventors determined that facilitating the grouping of segments from a plurality of different programs to be viewed within a single presentation defined by a single description scheme is beneficial and not previously possible. A "virtual program" consisting of segments from a plurality of different programs may be dynamically constructed and presented, without the need for physically creating the program on a persistent storage medium. Thus the description scheme syntax may
15 facilitate the identification of the "virtual program" where the relevant segments may be located with multiple different segments of multiple different programs being identified within a single description scheme syntax, as illustrated in FIG. 9. Thus for example, it is possible to view various news reports from different news sources in a sequence without
20 having to individually manually select each news report from each news source.

Principally, existing description schemes for audiovisual content permit linking to external content, such as a web site of a news article on a relevant topic or other material of related interest to the content. However, it was determined that more focused external content may be selected for the user if the content was associated with segments
25 of the video, as opposed to the entire video. In this manner, a single video may include multiple links to external content, each link being associated with a different portion of the video.

5 Another issue that arises with respect to selecting segments within multiple different video streams is the different techniques that may be used to indicate "time". For example, the frame rate is not always the same, with movie film typically being 24 frames per second and television being 30 frames per second (each frame consisting of two fields). In addition, the time base for MPEG-1, DVD's, MPEG-2, VCR, Movies, Internet
10 based streaming media, etc., are not the same. Accordingly, the modified description scheme includes a time base indicator of the time base associated with the particular segment, as illustrated in FIG. 10. With the time base defined, either explicitly or implicitly, the beginning of the segment is defined together with an offset indicating the duration of the segment. The offset is in time units, such as seconds, minutes, frames,
15 fields, etc., and may have a default or implicit definition.

The following may be used to describe the descriptive properties of segments:

```
<complexType name="BasicSegmentDescriptionType">
  <sequence>
    <element name="Title" type="mpeg7:TitleType"
      minOccurs="0" maxOccurs="unbounded"/>
    <element name="Synopsis" type="tva:SynopsisType"
      minOccurs="0" maxOccurs="unbounded"/>
    <element name="Keywords" type="mpeg7:KeywordAnnotationType"
      minOccurs="0" maxOccurs="1"/>
    <element name="RelatedMaterial" type="tva:RelatedMaterialType"
```

5 minOccurs="0" maxOccurs="unbounded"/>

</sequence>

</complexType>

The terms and definitions may be as follows:

	BasicSegmentDescriptionType	Defines basic segment description.
10	Title	A title of the segment (optional). A segment can have several titles when necessary, e.g. in different languages.
15	Synopsis	A synopsis of textual description of the segment (optional). A segment can have several synopses when necessary, e.g. in different languages or lengths.
	Keywords	A keyword associated with the segment (optional). A segment can have several keywords when necessary, e.g. in different languages.
20	RelatedMaterial	A link to external material related to the segment (optional). A segment can have multiple links.

The following element and complex type may be used to define a segment.

<element name="Segment Information"

type="tva:SegmentInformationType"/>

5 Description A description of the content of the segment.

10 SegmentLocator Locates the segment within a program (instance) in terms of
start time and duration (optional). Defined as an MPEG-7
datatype, MediaTimeType (See Sec. 6.3.9 of MPEG-7 for a
detailed description). If the duration is not specified, the
segment ends at the end of the program. If the timeBase
and/or timeUnit attributes for the SegmentLocator element are
instantiated, they override the default or global time base and
time unit definitions provided with the
SegmentInformationTable description.

15 KeyFrameLocator Locates a key frame of the segment within a program in terms
of a time point (optional). Defined as an MPEG-7 datatype,
MediaTimeType (See Sec. 6.3.9 of MPEG-7 for a detailed
description). MediaDuration and MediaIncrDuration elements
of a KeyFrameLocator element should not be used. Multiple
key frames may be associated with a single segment. If the
timeBase and/or timeUnit attributes for the KeyFrameLocator
element are instantiated, they override the default or global
time base and time unit definitions provided with the
SegmentInformationTable description.

20

5 segmentId The unique identifier of the segment.

version The version number of the segment information (optional).

The following element and complex types may be used to define segment grouping.

<element name="SegmentGroupInformation"

10 type="tva:SegmentGroupInformationType"/>

<complexType> name="SegmentGroupInformationType">

 <sequence>

 <element name="ProgramRef" type="tva:CRIDRefType"/>

 <element name="Description" type="tva:BasicSegmentDescriptionType"

15 minOccurs="0"/>

 <element name="GroupInterval" minOccurs="0">

 <complexType>

 <attribute name="ref" type=xsd:IDREF"/>

 </complexType>

20 </element>

 <choice minOccurs="0" maxOccurs="1">

 <element name="Segments">

 <complexType>

```
5          <attribute name="refList"
              type=xsd:IDREFS" use="required"/>
          </complexType>
        </element>
        <element name="Groups"
10          <complexType>
              <attribute name="refList"
                  type=xsd:IDREFS" use="required"/>
              </complexType>
            </element>
15    <choice>
</sequence>
        <element name="KeyFrameLocator" type="mpeg7:MediaTimeType"
            minOccurs="0" maxOccurs="unbounded"/>
        <attribute name="groupId" type="xsd:ID"
20            use="required"/>
        <attribute name="groupType" type="tva:SegmentGroupTypeType"
            use="required"/>
        <attribute name="ordered" type="xsd:boolean"
            use="optional"/>
25    <attribute name="numberOfSegments" type="xsd:positiveInteger"
        use="optional"/>
    <attribute name="numberOfKeyFrames" type="xsd:positiveInteger"
```

```

5          use="optional"/>

<attribute name="numberOfKeyFrames" type="xsd:positiveInteger"
          use="optional"/>

<attribute name="duration" type="mpeg7:mediaDurationType"
          use="optional"/>

10    <attribute name="topLevel" type="xsd:boolean"
          use="optional"/>

    <attribute name="version" type="xsd:integer"
          use="optional"/>

</complexType>

15 <simpleType name="SegmentGroupType">
    <list>
        <simpleType>
            <restriction base="xsd:string">
                <enumeration value="highlights"/>
20    <enumeration value="highlights/objects"/>
        <enumeration value="highlights/events"/>
        <enumeration value="bookmarks"/>
        <enumeration value="bookmarks/objects"/>
        <enumeration value="bookmarks/events"/>
25    <enumeration value="themeGroup"/>
        <enumeration value="preview"/>

```

```

5      <enumeration value="preview/title"/>

      <enumeration value="preview/slideshow"/>

      <enumeration value="tableOfContents"/>

      <enumeration value="synopsis"/>

      <enumeration value="shots"/>

10     <enumeration value="alternativeGroups"/>

      <enumeration value="other"/>

      </restriction>

      </simpleType>

      </list>

15  </simpleType>

```

The names and definitions may be as follows:

SegmentGroupInformation An element that instantiates the SegmentGroupInformationType.

SegmentGroupInformationType Defines an individual segment group.

20 **ProgramRef** A reference to the program this segment belongs to. When the member segments/groups are collected from different programs (e.g. to create a "virtual program") the ProgramRef element references a group CRID that is resolved into the individual programs CRIDs.

A description of the content of the segment group.

References a single segment that defines the temporal range of the segment group.

Defines the segments that are part of this group by providing a list of references to the identifiers of elements of type

SegmentInformationType (optional). The order of the references to segments in this list determines the ordering of segments within this group.

Defines the segment groups that are subgroups of this group by providing a list of references to the identifiers of elements of type `SegmentGroupInformationType` (optional). The order of the references to segment groups in this list determines their ordering within this group.

SegmentGroupInformationType (optional). The order of the references to segment groups in this list determines their ordering within this group.

Locates a key frame of the segment group within a program in terms of a time point (optional). Defined as an MPEG-7 datatype, `MediaTimeType` (See Sec. 6.3.9 of MPEG-7 for a detailed description). `MediaDuration` and `MediaIncrDuration` elements of a `KeyFrameLocator` element should not be used.

datatype, `MediaTimeType` (See Sec. 6.3.9 of MPEG-7 for a detailed description). `MediaDuration` and `MediaIncrDuration` elements of a `KeyFrameLocator` element should not be used.

[illegible]

groupID

groupType

10

Specifies whether the given segment group presents an ordered list (i.e. whether order of the segment or segment groups within the given segment group is significant) (optional).

15

The number of segments in the segment group (optional). The value of this attribute specifies only the segments that are direct members of the segment group.

The number of key frames in the segment group (optional). The value of this attribute specifies only the key frames of the segments that are direct members of the segment group.

20

The sum of the durations of the segments contained within this group (optional). This duration corresponds to the sum of the

5 durations of the segments that are direct members of the segment group.

topLevel Specifies whether the given segment group is a top-level group (optional).

version The version number of the segment group (optional).

10 SegmentGroupTypeType A simple type that specifies the valid types of segment groups.

The allowed types may be defined as follows:

highlights - The group of segments represents selected highlights from one or more programs. The segments that constitute a highlight group are presented seamlessly in sequence, without any temporal interruptions.

highlights/objects - The group of segments represents selected highlights from a program (or programs) that share a common object or objects (e.g. Seinfeld highlights with Kramer).

5

highlights/events - The group of segments represents selected highlights from a program (or programs) that share a common event or events (e.g. touchdowns in the Super Bowl).

10

bookmarks - The segment group defines a set of access points to a program. If the member segments of a segment group of type bookmarks contain segment duration information, this duration information shall be ignored, and the segments shall be treated as "open-ended".

15

a

bookmarks/objects - The segment group defines a set of access points to a program, where the selected access points share a common object or objects. If the member segments of a segment group of type bookmarks/objects contain segment duration information, this duration information shall be ignored, and the segments shall be treated as "open-ended."

20

bookmarks/events - The segment group defines a set of access points to a program, where the selected access points share a common event or events. If the member segments of a segment group of type

5 bookmarks/events contain segment duration information, this duration information shall be ignored, and the segments shall be treated as "open-ended."

10 themeGroup - The segment group comprises segments that share a common topic or theme. The common theme can be specified in the segment group description. A theme group does not necessarily require direct continuous playback.

preview - The segment group defines a preview of a program.

preview/title - The segment group defines a preview of a program, where the preview serves as a promotional title or trailer for the program.

15 preview/slideshow - The segment group defines a preview of a program, where the preview serves as a compact slideshow of the program content.

tableOfContents - The segment group defines a navigable table of contents for the program.

synopsis - The segment group provides a summary or synopsis of the program.

20 shots - The segment group provides a list of the shots in the program.

5

alternativeGroups - Each member of this type of segment group provides an alternative view of representation, with the same functionality but different durations or levels of detail.

other - The functionality of the segment group does not fit into any of the other types defined.

10

Various validity constraints may be imposed on the proposed description scheme to ensure that (i) it fits the data model of FIG. 11, and (ii) the sequence and relationships of the various segments and segment groups are unambiguously defined. These constraints, which are implicit in the description schemes, are outlined below for clarity:

- (a) A segment group may contain either segments, or subgroups, but not both.
- (b) A segment group of type "alternativeGroups" may not contain segments and shall only contain subgroups.
- (c) A segment group of any type other than "tableOfContents" and "alternativeGroups" may only contain segments. A group of type "tableOfContents" may contain other segment groups of type "tableOfContents".

15

20

These validity constraints reduce the complexity of the resulting descriptions by limiting the degree of nesting in the hierarchy. The navigation order of segments or segment groups is determined by the order of references to the segments in a segment group.

The entities may be defined as follows:

5	Program	The program entity represents an editorially coherent audiovisual content.
	Program Location	Program Location provides a physical location where the program is available. A program may be available at multiple program locations; however, the time lines of the different instances of a
10		program are preferably identical.
	Segment	A segment is a continuous fragment of a program. A particular segment should belong to a single program, but it may be a member of multiple segment groups.
	Segment Group	A segment group is a collection of segments that are grouped together, for a particular purpose or due to a shared property. A
15		segment group may contain segments, or other segment groups.
The entity relationships may be defined, as follows.		
	Program-to-Segment	A Segment is part of a single program. A Program may contain multiple segments.
20	Segment-to-Segment Group	A Segment may belong to zero or more Segment Groups. A Segment Group may contain zero or more Segments (possibly from multiple Programs).
	Segment Group-to-Segment Group	A Segment Group may be a member of zero or more Segment Groups, and it may contain zero or more Segment
25		Groups. A Segment Group may contain either segments, or subgroups, but not both.

5 The following element and complex type define a structure for holding
segmentation-related metadata.

```
<element name="SegmentInformationTable"
```

```
  type="tva:SegmentInformationTableType"/>
```

```
<complexType name="SegmentInformationTableType">
```

10 <sequence>

```
    <element name="SegmentList">
```

```
      <complexType>
```

```
        <sequence>
```

```
          <element ref="tva:SegmentInformation"
```

15 minOccurs="0" maxOccurs="unbounded"/>

```
        </sequence>
```

```
      </complexType>
```

```
    </element>
```

```
<element name="SegmentGroupList"
```

20 <complexType>

```
    <sequence>
```

```
      <element ref="tva:SegmentGroupInformation"
```

```
        minOccurs="0" maxOccurs="unbounded"/>
```

```
    </sequence>
```

25 </complexType>


```

5      </element>

      <element name="TimeBase Reference" minOccurs="0" maxOccurs="unbounded"/>

        <complexType>

          <choice>

            <element name="RefMediaTime" type="mpeg7: mediaTimePointType"/>

10          <element name="RefURI" type="xsd:anyURI"/>

          </choice>

        </complexType>

      </element>

    </sequence>

15    <attribute name="timeUnit" type="mpeg7:mediaDurationType"

      use="default" value="PT1N1000F"/>

    <attribute name="version" type="xsd:integer" use="optional"/>

  </complexType>

```

The names and definitions may be as follows:

20	SegmentInformationTable	An element that instantiates the SegmentInformationTableType.
	SegmentInformationTableType	Defines a structure for holding segmentation-related metadata.

5 timeUnit Specifies the duration of the time intervals used in the incremental specifications of relative time points and duration. Default time unit is milliseconds.

version The version number of the segment information table (optional).

10 The terms and expressions employed in the foregoing specification are used therein as terms of description and not of limitation, and there is no intention in the use of such terms and expressions of excluding equivalents of the features shown and described or portions thereof, it being recognized that the scope of the invention is defined and limited only by the claims that follow.

10000000.000000